PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
G06K 9/64, G06T 7/00

A1

(11) International Publication Number: WO 96/39678

(43) International Publication Date: 12 December 1996 (12.12.96)

(21) International Application Number: PCT/US96/07859

(22) International Filing Date: 29 May 1996 (29.05.96)

(30) Priority Data:
462,352 5 June 1995 (05.06.95) US

(71) Applicant: SHELL OIL COMPANY [US/US]; Intellectual Property, P.O. Box 2463, 900 Louisiana, Houston, TX 77252-2463 (US).

(72) Inventor: BEATTY, David, Andrew; 3980 Carrigan Court #405, Burnaby, British Columbia V3N 4S6 (CA).

(74) Agent: CHRISTENSEN, Del, S.; Shell Oil Company, Intellectual Property, P.O. Box 2463, 900 Louisiana Street, Houston, TX 77252-2463 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: OBJECT RECOGNITION METHOD

(57) Abstract

A method to identify the location and orientation of an article is provided, the method comprising the steps of: obtaining an image of the article with the article having a known orientation and location relative to a camera; creating a X and Y template edge matrix from the image of the article; creating a plurality of sets of modified template edge matrices, each of the sets of modified template edge matrices being an X and Y template edge matrix with the article in a different orientation; capturing a digital visual image containing the article, the digital image being a matrix of pixels; creating X and Y article edge matrices from the matrix of pixels; quantifying difference between each of the sets of modified template edge matrices and the X and Y article edge matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices; and identifying the location and orientation of the article as the orientation of the article represented by the set of modified template edge matrices at the location within the bounds of the X and Y article edge matrices with the minimal quantified differences between the modified template edge matrices and the X and Y article edge matrices.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

DESCRIPTION

VISUAL RECOGNITION METHOD

Technical Field

This invention relates to a method for identifying the location and orientation of a known article within a visual field.

Background Art

U.S. Pat. No. 5,379,353 suggests a differential analysis circuit that utilizes a step to identify edge vectors for identification of such things as a road for a mobile robot. A digital image captured from a video camera is processed using an algorithm that includes generation of a differential of brightness along each row of pixels, and presumably also along each column of pixels. The absolute value of the differential brightness represents a change in the picture, and a differential that exceeds a threshold is identified as a possible edge to a road.

U.S. Pat. No. 5,381,155 suggests a speed detection system that identifies moving vehicles in the view of a fixed camera, measures the speed at which the vehicles are moving, and identifies a license plate number from the vehicle. Commercially available systems are disclosed that are said to be capable of identifying the license plate within a captured image and then reads the numbers and letters within the license plate number.

U.S. Pat. No. 5,381,489 suggests a system for recognition of characters on a medium. This system includes making a window of a possible character from the medium, and then comparing that window to each template within a set. The entire set of templates must be screened each time a character is identified. The templates are generated based on previously recognized characters from the document where the initial recognition requires a more rigorous comparison to different character features.

A problem faced in visual recognition is to recognize the location, within the view of a camera, and the orientation, of a particular article where the article may

be one of a relatively few possible articles already identified from a library of potential articles. The possibility of a variety of lighting conditions and shadows make such recognition difficult. There are also typically constraints on the amount of computer data storage available at the site of a desired visual recognition facility. Therefore, templates of different orientations an scales of the different articles can generally not be generated and stored initially.

Such a problem in visual recognition is encountered when visual recognition is used as a means to identify vehicles or determine the orientation of vehicles in an automated refuelling system. For example, in U.S. Pat. No. 3,527,268 it is suggested that vehicle identification in an automated refuelling system can be achieved in a fully automated method by a photo-electric means to detect the silhouette of the automobile. How this is to be done is not suggested.

It is therefore an object of the present invention to provide a method to identifying the location and orientation of an article, wherein the method is capable of identifying the location and orientation of the article in a variety of natural and artificial lighting conditions, and wherein a large number of templates do not have to be digitally stored.

Disclosure of the Invention

These and other objects of the invention are accomplished by a method to identify the location and orientation of an article, the method comprising the steps of: obtaining an image of the article with the article having a known orientation and location relative to a camera; creating a X and Y template edge matrix from the image of the article; creating a plurality of sets of modified template edge matrices, each of the sets of modified template edge matrices being a X and Y template edge matrix with the article in a different orientation; capturing an digital visual image containing the article,

the digital image being a matrix of pixels; creating X and Y article edge matrices from the matrix of pixels; quantifying difference between each of the sets of modified template edge matrices and the X and Y article edge matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices; and identifying the location and orientation of the article as the orientation of the article represented by the set of modified template edge matrices at the location within the bounds of the X and Y article edge matrices with the minimal quantified differences between the modified template edge matrices and the X and Y article edge matrices.

This method can be readily adapted to identification of a location and orientation of a vehicle within a bay for automated refuelling purposes. The make and model of the vehicle can be identified by another means, such as for example, driver manual input, a magnetic or optical strip, or a passive or active transponder located on the vehicle. With the make and model identified, or limited to one of a small number of possibilities (such as when more than one transponder signal is being received), base templates can be retrieved from storage. The base templates can be prepared from a digital visual image of the known make and model of vehicle with the vehicle positioned at a known location with respect to the camera, and the image processed by generation of X and Y edge matrices. After the make and model of the vehicle are identified, a series of modified templates are created from the retrieved templates by rotation of the template edge matrices to different angles from the initial orientation and/or scaling the matrices to represent different distances from the camera. Thus, only one set of base templates (or one set of X and Y edge matrix templates) needs to be stored in the data base for each vehicle. A captured visual image containing the vehicle within the refuelling facility can then be processed to generate article edge matrices, and compared to the modified

templates, with each modified template being compared to the article edge matrices at different locations within the article edge matrices.

Preferably, a mask of the template is prepared so that only the outlines and/or internal edges of the article, and not the surrounding area, is compared to the actual article edge matrices. The mask also provides expected dimensions of the article so that only locations within the article edge matrices within which the article would fit would be searched for the article, and the article can be identified with a position relatively close to the edge of the view.

Separately comparing the X and Y edge template matrices with the article X and Y edge matrices significantly improves the robustness of the method, and results in reliable fits being found quickly in a variety of light conditions, with partial obstruction of the view of the article, and with partial masking by dirt, leaves, grass, and other articles that may be present in a relatively uncontrolled environment.

Detailed Description of the Invention

A camera is typically used in the practice of the present invention to capture a visual image of an article in a known position and orientation. A digital image can be captured using one of the commercially available framegrabber hardware and associated software packages. The digital image is a matrix of pixels, each of the pixels having a number that corresponds linearly to a brightness. A color image can be utilized, in which case the image is represented by three matrices, one each for red, green and blue. Typically, the images of about 256 by 240 pixels are preferred for the practice of the present invention because such a number of pixels results in sufficient resolution and is within the capacity of relatively inexpensive video cameras. The video camera may generate an image of about twice the resolution of a 256 by 240 matrix, in which case the image can be reduced by averaging adjacent pixels to

create a matrix of pixels having one half the height and one half the width by averaging blocks of four pixels.

Signal to noise ratios can be increased by averaging two or more consecutive images.

Edge matrices may be generated from both the images containing the article in a known location and orientation relative to the camera, and the images containing the article within which the location and orientation of the articles are to be determined, by applying operators such as the following:

$$\frac{1}{8}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

and;

$$\frac{1}{8}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

As can be seen from these operators, they each result in a matrix in which the elements will sum to zero. The absolute values of the elements of the resulting matrices indicate the change in brightness along the x and y axis respectively. The use of edge magnitudes helps make the appearance invariant to direction of light and color of the object. Producing these edge matrices therefore results in images that can be compared with templates inspite of significant differences in lighting or color of the article (although color could be identified as well in the practice of the present invention). The results of these two operators can be summed to obtain one edge vector matrix, but in the practice of the present invention, it is significant that the two are not combined for comparison. Not combining the two greatly increases the robustness of the algorithm, i.e. the ability of the algorithm to identify outlines when the articles are masked with dirt, partially obscured, or subjected to varying light conditions.

A mask is preferably place over the image of the edge matrices of the article with the know location and orientation so that only the know outlines of the article are considered. The masked regions can be referred to as "don't care" regions, because edge data in these regions will be ignored when fitting the edge matrices of the article in the known position to the edge matrices of the article within which the location and orientation of the article is to be identified.

The dimensions of the mask can also define limits of the locations within the image containing the article that could contain an image of the article. For example, if the mask were fifty pixels by fifty pixels within an image of 256 pixels by 240 pixels, then only locations within the middle 206 by 190 pixels could be the center of the mask, if the entire article is within the image.

When a color image is used, edge matrices are preferably generated for each color, and then the three edge matrices are preferably combined to form one X or Y edge matrix. This combination can be by summing the absolute values of the three edge matrices (and dividing the sum by three), by selecting the maximum value of the edge matrix among the three, by calculating an average, or by taking the square root of the sum of the squares of corresponding elements of the edge matrices. It may also be possible to consider two of the three colors in, for example, one of the preceding ways. The use of a color image improves the fit to a template by providing that more edge information is extracted from an image. For example, an edge image at an interface between colors can be identified even if the interface is between surfaces having similar brightness. Use of color images increases the cost of the camera, and the increases amount of data processing required to take advantage of having three sets of images, but is preferred if the difficulty of the application warrants the additional expense.

The image of the article may be reduced by, for example averaging adjacent pixels. An image is therefore created that contains fewer pixels for comparison of the modified templates to the article edge matrices for finding an initial position and orientation estimate. For example, a matrix of 256 by 240 pixels could be reduced to a matrix of 64 by 60 pixels by two successive averaging and subsampling operations. Comparisons of the reduced matrices can be accomplished much more quickly, and then the fit of the reduced matrices can be used as a starting point for finding a fit for the larger matrices. Generally, only locations within a few (one to three) pixels of the pixels averaged into the best fit result of the reduced matrix need to be compared at higher levels of resolution.

Reducing the matrices can significantly reduce computing time required to compare the template edge matrices with the article edge matrices. Two reductions, each being two for one linear reductions, are recommended. Thus, each reduction therefore reduces the amount of information to be considered by a factor of four. The combined reductions reduce the amount of information by a factor of sixteen. Further, each of the parameters for which templates are prepared are of lower resolution, resulting in fewer sets of rotated templates, and at fewer locations within the view of the article edge matrix. Initial searches within the reduced matrixes can therefore be performed in two or three orders of magnitude less time than if the article edge matrix was searched at the original level.

When the article which the image is to be searched for is know, a template of that article, preferably as two edge matrices with a mask, can be selected from a data base. The template can then be modified to represent the article in a plurality of orientations. By orientations, it is meant that the two dimensional image of the article is rotated about the axis of the view of the camera, rotated to an angled view of the object, and/or scaled to represent

changes in distance from the camera. Increments of, for example, two to three degrees of rotation can be used to obtain a sufficient number of orientations that one should have a clearly best fit for the article in an particular orientation.

For an application such as an overhead camera identifying a vehicle's position within a bay of an automated refuelling system, an expected orientation can be predicted (most drivers drive in relatively straight), and it can be also predicted that the actual orientation will not be more than a certain variation (for example, plus or minus twenty degrees) from the expected orientation. Thus, only a limited number of modified template edge matrices need to be created. But creating these modified templates after the vehicle make and model have been identified considerably reduces the amount of computer storage needed to store template matrices.

Rotation of the template matrices about an axis essentially normal to a plane of the two dimensional view of the video camera (or "transforming" the image to the new orientation) is readily accomplished by well known methods. Such transformations are preferably performed by calculating the point within the original matrix of pixels each pixel within the transformed matrix would lie so that the four pixels of the original matrix surrounding the center of the pixel from the transformed matrix can be used to interpolate a value for the pixel of the transformed matrix. Again, methods to accomplish these interpolations are well known.

The templates could also be created with the image scaled to represent the article located at different distances from the camera. This scaling is accomplished by changing the dimensions from a center point of the camera view inversely proportional with the distance from the camera. This scaling is preferably performed based on distances from an effective pinhole, where the effective pinhole is defined as a point through which a perspective

projection is made by the camera. This effective pinhole would therefore be slightly behind the lens of the camera.

A more difficult problem is to identify a location and orientation of a known article or outline when viewed at an angle significantly different from normal to a plane containing the article or outline. For example, a camera located on a refuelling apparatus may need to locate a gasoline nozzle cover lid from an position that does not allow viewing of the cover lid with the camera facing perpendicular to the plane of the cover lid. A rectangular lid cover would therefore not appear to the camera to be rectangular. The distortion from a rectangular shape would depend upon both the angle and the relative position of the lid with respect to the centerline of the camera's view (known as the optical axis).

Geometric distortion can be eliminated from images that are not normal to the optical axis if the article of the image can be approximated by a planar image. If the angle of the optical axis from perpendicular of the planar image of the article is known (i.e., the image to be searched for the article), geometric distortion can be removed, and images obtained that represent a transformation to perpendicular views of the article in the image to be searched. Likewise, if the templates are created wherein the optical axis is not perpendicular to the plane of the template, geometric distortion can be removed from the templates by such a transformation. If the angle from normal to the optical axis is not known for the image to be searched, this angle can be another search parameter.

Such transformation to a perpendicular view is simplified by the fact that the transformation is the same for a given angle between the optical axis and the normal of the plane approximately containing the article, regardless of the displacement between the camera and the plane, provided that the target remains in the limits of the view of the camera.

The preferred transformation method in the practice of the present invention, rather than to place a pixel from an article image matrix within a transformed image, will take a pixel location from the transformed matrix and calculate the location of that pixel within the article image matrix. An interpolation is then performed using four pixel values of the article image matrix surrounding the position of the inversely transformed matrix pixel to determine the value of the pixel in the transformed image. The following equations provide the location of a pixel from the transformed image on the article image, for the case of the article plane normal being perpendicular to the image X axis:

$$\rho_x=\frac{\sigma_x\left(\frac{P_o'}{P_o}\right)}{b+a\left(\frac{P_{zc}}{P_o}\right)-a\left(\frac{P_o'}{P_o}\right)\sigma_y} \qquad (3)$$

and:

$$\rho_y=\frac{a-b\left(\frac{P_{zc}}{P_o}\right)+b\left(\frac{P_o'}{P_o}\right)\sigma_y}{b+a\left(\frac{P_{zc}}{P_o}\right)-a\left(\frac{P_o'}{P_o}\right)\sigma_y} \qquad (4)$$

where:

$$a=\sin(\theta) \qquad (5)$$

and:

$$b=\cos(\theta) \qquad (6)$$

and: ρ_x is the ratio of actual article image plane x position to P_O,

ρ_y is the ratio of actual article image plane y position to P_O,

P_O is the perpendicular distance from the effective pinhole to the actual article,

P_o' is the distance from the plane of the transformed image to the effective pinhole,

P_{zc} is the vertical displacement of the camera of the transformed image relative to the camera position in the actual image,

σ_y is the y coordinate value in the transformed image,

σ_x is the w coordinate value in the transformed image, and

θ is the downward pitch angle of the plane normal to the camera.

For θ of up to about fifty degrees, the following ratios can be used to fit a good portion of the original image into the transformed image:

$$\left(\frac{P_o'}{P_o}\right) = [(b+a\rho_{y1})(b-a\rho_{y1})]^{-1} \qquad (7)$$

and

$$\left(\frac{P_{zc}}{P_o}\right) = \frac{ab(1+\rho_{y1}^2)}{(b+a\rho_{y1})(b-a\rho_{y1})} \qquad (8)$$

where ρ_{y1} is half of the vertical height of the original image.

Although modified templates can be created with rotations and changes in distances from the camera, a plurality of such rotations and changes could result in an exceedingly large number of modified templates. It is therefore preferred that searches are carried out over one variable out of the possible rotating, scaling, and angled views in the practice of the present invention.

If the orientation of the article with respect to rotation within a plane perpendicular to the camera view is expected to be within about twenty degrees of the orientation of the article having the known orientation, the template X and Y edge images may be simply individually rotated to form the modified edge template images prior to comparing the modified template edge images to the article edge images. When more than about twenty degrees of

rotation is possible, a new set of edge images is preferably created based on a combination of the original edge images. The X and the Y edge image values together represent an edge vector having an angle (arctan(Y/X)) and a magnitude $((X^2+Y^2)^{1/2})$. This angle may be rotated by the angle of rotation of the template and new X and Y components calculated. Typically, only the absolute values of the X and Y components are stored, and therefore edge vectors in the first or third quadrant must be differentiated from edge vectors of the second or fourth quadrant. Edge vectors in the third and fourth quadrants could be considered as their negative vectors in the first and second quadrants respectively, and therefore just two quadrants of vectors need be identified. Quadrants of edge vectors can be identified with a single additional binary template generated from the original template image, the binary template having pixels representing whether the edge magnitude vector at that point represents an edge whose direction vector is in the first or third quadrant, or the second or fourth quadrant. This template can be automatically generated from the template image. This requires very little additional storage space, and can be used during a rotation operation to adjust the X and Y edge magnitude weights to their exact proper proportion at very little extra computational cost. Rotation of the edge matrices by any amount of rotation can thereby be made completely valid.

The following equation is convenient for the purpose of quantifying the differences between the modified template edge matrices and the article edge matrices because commercially available image processing cards are available to quickly generate the comparisons:

$$\rho(x,y) = \frac{\sum_{ij} X_{ij} X'_{(x+i)(y+j)} + \sum_{ij} Y_{ij} Y'_{(x+i)(y+j)}}{\sqrt{\sum_{ij} (X_{ij})^2 + \sum_{ij} (Y_{ij})^2} \sqrt{\sum_{ij} (X'_{(x+i)(y+j)})^2 + \sum_{ij} (Y'_{(x+i)(y+j)})^2}} \tag{9}$$

where X is a X template edge matrix of i by j pixels rotated to an orientation to be tested against a portion of the image matrix,

Y is a Y template edge matrix of i by j pixels rotated to an orientation to be tested against a portion of the image matrix,

X' is a portion of an image X edge matrix of i by j pixels located at a position of coordinates x,y on the X image edge vector matrix,

Y' is a portion of an image Y edge matrix of i by j pixels located at a position of coordinates x,y on the Y image edge vector matrix, and

$\rho(x,y)$ is a grey scale edge correlation normalized for point (x,y).

The grey scale edge correlation will be a number between zero and one, with one being a perfect match. Grey scale correlations are performed for each x and y within the article edge matrix for which the entire modified template edge matrix can fit within the article edge matrix. The resulting grey scale correlation that is the closest to approach unity is the closest fit. Interpolation between variables can be achieved using linear or squared weighing above a noise threshold. Such variables may be, for example, angle of rotation, or x and y locations.

Portions of the calculations to generate these grey scale edge matrices can be quickly made using a GPB-1 auxiliary card-AlignCard.

"Don't care" regions may also fall within the boundaries of the i by j dimensioned matrices of the modified template edge matrix. Pixels in the template identified as "don't care" are preferably not used in the summations of the terms of Equation 9.

Because the grey scale edge matrix correlation result is very sensitive to relative displacement of an object's template and test image, a smoothing operation may be performed prior to comparison of the two. Although reducing the matrices as described above has a smoothing effect, a further smoothing operation may also be included. This smoothing operation may be performed on each before the correlation is calculated, but after the subsampling to a current search level. A preferred smoothing operation is a Gaussian approximation, given by the following convolution kernel:

$$\frac{1}{8}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad (10)$$

When this smoothing is applied, it is preferably applied to both the article edge matrix and the modified template edge matrix.

A preferred application of the method of the present invention is an automated refuelling methods disclosed in U.S. Pat. Appl. Nos. 461,280 (Docket No. TH0622), 461,281 (Docket No. TH0572), and 461,276 (Docket No. TH0573, the disclosures of which are incorporated herein by reference.

The embodiments described above are exemplary, and reference is made to the following claims to determine the scope of the present invention.

CLAIMS

1. A method to identify the location and orientation of an article, the method comprising the steps of:

obtaining an image of the article with the article having a known orientation and location relative to a camera;

creating a X and Y template edge matrix from the image of the article;

creating a plurality of sets of modified template edge matrices, each of the sets of modified template edge matrices being a X and Y template edge matrix with the article in a different orientation;

capturing an digital visual image containing the article, the digital image being a matrix of pixels;

creating X and Y article edge matrices from the matrix of pixels;

quantifying difference between each of the sets of modified template edge matrices and the X and Y article edge matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices; and

identifying the location and orientation of the article as the orientation of the article represented by the set of modified template edge matrices at the location within the bounds of the X and Y article edge matrices with the minimal quantified differences between the modified template edge matrices and the X and Y article edge matrices.

2. The method of Claim 1 further comprising the steps of

smoothing the modified template edge matrices and the article edge matrices by averaging adjacent pixels and subsampling to obtain reduced matrices with a reduced number of pixels;

quantifying the difference between each of the reduced modified template edge matrices and the reduced article edge matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices; and

quantifying the difference between each of the sets of modified template edge matrices and the X and Y article edge matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices only for locations near the location and the orientation having minimum differences between the reduced matrices.

3. The method of Claim 1 wherein the edge matrices are obtained by applying to the image to obtain horizontal and vertical edge matrices respectively the operators:

$$\frac{1}{8}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

and;

$$\frac{1}{8}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

4. The method of Claim 1 wherein a plurality of modified template edge matrices are created with the template edge matrix scaled to represent different distances from the camera.

5. The method of Claim 1 wherein a plurality of modified template edge matrices are created with the template edge matrix transformed to represent different angles from normal to the optical axis.

6. The method of Claim 1 wherein portions of the template edge matrices outside of outlines of the article are ignored when quantifying the difference between each of the modified template edge matrices and the article edge matrices.

7. The method of Claim 1 wherein the plurality of locations is every contiguous set of pixels within the article edge matrices within which a modified template edge matrices will fit.

8. The method of Claim 2 wherein the plurality of locations for which the difference between each of the

reduced modified template edge matrices and the reduced article edge matrices are quantified include every contiguous set of pixels within the reduced modified template edge matrices will fit.

9. The method of Claim 1 wherein a color image is obtained, and single color edge matrices are created for more than one color, and then combined to obtain the template edge matrices.

10. The method of Claim 9 wherein three set of single color edge matrices are created.

11. The method of Claim 2 wherein edge matrices are obtained by applying to the image to obtain vertical and horizontal edge matrices respectively the operators:

$$\frac{1}{8}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

and;

$$\frac{1}{8}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

12. The method of Claim 11 wherein portions of the template edge matrices outside of outlines of the article are ignored when quantifying the difference between each of the modified template edge matrices and the article edge matrices.

13. The method of Claim 12 wherein the plurality of locations comprises every contiguous set of pixels within the matrix of pixels within which a modified template edge matrices will fit.

14. The method of Claim 12 wherein the plurality of locations for which the difference between each of the reduced modified template edge matrices and the reduced article edge matrices are quantified comprise every contiguous set of pixels within the reduced modified template edge matrices will fit.

15. The method of Claim 14 wherein a color image is obtained, and single color edge matrices are created for more than one color, and then combined to obtain the template edge matrix.

16. The method of Claim 15 wherein three sets of single color edge matrices are created and then combined to obtain the template edge matrix.

17. A method to identify the location and orientation of an article, the method comprising the steps of:

obtaining an image of the article with the article having a known orientation and location and distance from a camera;

creating a X and Y template edge matrix from the image of the article;

creating a plurality of sets of modified template edge matrices, each of the sets of modified template edge matrices being a X and Y template edge vector matrix with the article in a different orientation;

capturing an digital visual image containing the article, the digital image being a matrix of pixels;

creating X and Y article edge matrices from the matrix of pixels;

smoothing the modified template edge matrices and the article edge matrices by averaging adjacent pixels and subsampling to obtain reduced matrices with a reduced number of pixels;

quantifying the difference between each of the reduced modified template edge matrices and the reduced article edge matrices with the reduced modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices; and

quantifying the difference between each of the sets of modified template edge matrices and the article matrices with the modified template edge matrices placed at a plurality of locations within the bounds of the article edge matrices only for locations near the location and the

orientation having minimum differences between the reduced matrices; and

identifying the location and orientation of the article as the orientation of the article represented by the set of modified template edge matrices at the location within the bounds of the X and Y article edge matrices with the minimal quantified differences between the modified template edge matrices and the X and Y article edge matrices,

wherein portions of the template edge matrices outside of outlines of the article are ignored when quantifying the difference between each of the modified template edge matrices and the article edge matrices.

18. The method of Claim 17 wherein the plurality of locations for which the difference between each of the reduced modified template edge matrices and the reduced article edge matrices are quantified include every contiguous set of pixels within the reduced modified template edge matrices will fit.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/07859

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06K9/64 G06T7/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G06K G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	IEICE TRANSACTIONS, vol. E74, no. 6, 1 June 1991, pages 1728-1734, XP000262328 MASAKI I: "INDUSTRIAL VISION SYSTEMS BASED ON APPLICATION-SPECIFIC IC CHIPS" see chapters 2, 4, 7, 8; figures 1, 8, 9 ---	1-18
A	US,A,4 658 246 (KUPERMAN GILBERT G) 14 April 1987 see abstract ---	1,17

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
9 September 1996	17.09.96
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Sonius, M

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	COMPUTER VISION, GRAPHICS AND IMAGE PROCESSING, vol. 27, no. 1, July 1984, MA US, pages 97-114, XP002012880 FRANK P. KUHL: "Global Shape Recognition of 3-D Objects Using a Differential Library Storage" See introduction; figures 1, 2 -----	5

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

I ational Application No
PCT/US 96/07859

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A-4658246	14-04-87	NONE	

Form PCT/ISA/210 (patent family annex) (July 1992)

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 043 689 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
11.10.2000 Bulletin 2000/41

(51) Int Cl.7: G06T 7/00

(21) Application number: 00303009.5

(22) Date of filing: 10.04.2000

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 08.04.1999 JP 10188599

(71) Applicant: FANUC LTD
Minamitsuru-gun, Yamanashi 401-0597 (JP)

(72) Inventors:
- Watanabe, Atsushi
Suginami-ku, Tokyo 168-0072 (JP)
- Arimatsu, Taro
Minamitsuru-gun, Yamanashi 401-0310 (JP)

(74) Representative: Billington, Lawrence Emlyn
Haseltine Lake & Co.
Imperial House
15-19 Kingsway
London WC2B 6UD (GB)

(54) Image processing apparatus

(57) An image processing apparatus is capable of detecting position and posture of individual workpieces randomly arranged in a pile and having identical shapes. Reference models are created from two-dimensional images of a reference workpiece captured in a plurality of directions by a camera and stored. Also, the relative positions/postures of the workpiece with respect to the camera at the respective image capturing are stored. An image of a pile of workpieces is captured by the camera to obtain a two-dimensional image and the position/posture of the camera at the image capturing is stored. An image of a workpiece matched with one reference model is selected by matching processing of the reference model with the captured image. A three-dimensional position/posture of the workpiece with respect to the camera is obtained from the image of the selected workpiece, the selected reference model and position/posture information associated with the reference model. A picking-up operation for picking out a respective workpiece from a randomly arranged pile can be performed by a robot, based on the position/posture of the workpiece.

EP 1 043 689 A2

Description

[0001] The present invention relates to an image processing apparatus for detecting three-dimensional position and posture (orientation) of an object, and in particular to an image processing apparatus suitable for use in a bin-picking operation for taking out a workpiece one by one from a pile of workpieces using an industrial machine such as a robot.

[0002] The operation of taking out an individual workplace from a randomly arranged pile of workpieces or an aggregation of workpieces contained in a container of a predetermined size, which workpieces have identical shapes and different three-dimensional positions/postures, has been performed manually. In storing workpieces in a pallet or placing workpieces at a predetermined position in a machine or a device using a (dedicated) robot, since it has been impossible to directly take out individual workpieces one by one from the randomly arranged pile of workpieces by the dedicated robot, it has been necessary to rearrange the workpieces in advance so as to be picked out by the robot. In this rearrangement operation, it has been necessary to take out an individual workpiece from the pile manually.

[0003] The reason why individual workpieces having identical shapes and different three-dimensional positions/postures cannot be picked out by a robot from a randomly arranged pile of workpieces or an aggregation of workpieces contained in a container is that the position/posture of individual workplaces in the pile or the aggregation cannot be recognised, so that a robot hand cannot be placed to a suitable position/posture at which the robot hand can hold the individual workpiece.

[0004] An object of the present invention is to provide an image processing apparatus capable of detecting three-dimensional position and posture of individual objects in a randomly arranged pile or an aggregation in a container of a predetermined region, which have identical shapes and different three-dimensional positions/postures.

[0005] An image processing apparatus of the present invention comprises an image capturing device; and a memory storage reference model based on image data of a reference object captured by the image capturing device in a plurality of directions, and storing information of the capturing directions to be respectively associated with the reference models. The reference object may be an object for detection itself or an object having a shape identical to that of the object for detection.

[0006] The image processing apparatus also comprises a processor for performing matching processing of image data containing an image of the object for detection captured by the image capturing device with image data for reference models to select an image of an object matched with one of the reference models, and to obtain posture, or posture and position, of the object based on the selected image of the object, said one reference model and the information of the direction associated with said one reference model.

[0007] The reference models may be a part of the image data of the reference object or obtained by processing the image data of the reference object.

[0008] The image capturing device may be a camera for capturing two-dimensional image data, and in this case the image data of the reference object are captured by the image capturing device from a predetermined distance. Alternatively, the image capturing device may be a visual sensor for capturing three-dimensional image data, and when the three-dimensional visual sensor is adopted the image data containing an image of the object of detection may be two-dimensional arrangement data including distance information from the object of detection to the image capturing device, a part of said two dimensional arrangement data or a set of distance data.

[0009] The image capturing device may be attached to a wrist of a robot. Further, the image data of the reference object can be captured in a place different from a place where the detection of the object is performed, and supplied to the image processing apparatus on line or off line.

[0010] For a better understanding of the invention, and to show how the same may be carried into effect, reference will now be made, by way of example, to the accompanying drawings, in which:-

FIG. 1 is a diagram for showing a picking-up operation by a robot to take out an individual workplace from a pile of workpieces using an image processing apparatus according to an embodiment of the present invention;

FIGS. 2a-2d show examples of reference models;

FIG. 3 is a block diagram of a principal part of a robot controller;

FIG. 4 is a block diagram of the image processing apparatus according to an embodiment of the present invention;

FIG. 5 is a flowchart of the processing for creating reference models;

FIG. 6 is a flowchart of the processing for the picking-up operation;

FIG. 7 is a diagram showing an example of scanning motion of a visual sensor capable of obtaining distance data;

FIG. 8 is a diagram of the two-dimensional arrangement data containing distance data as image data obtained by the visual sensor;

FIG. 9 is a flowchart of processing for obtaining the two-dimensional arrangement data.

[0011] An embodiment in which an image processing apparatus of the present invention is used in combination with a robot system will be described. In this embodiment, an image of a pile of workpieces, which are objects for detection having identical shapes and randomly arranged as shown in FIG. 1, is captured by an

image capturing device (camera or visual sensor) 20, which is attached to a wrist of a robot RB and position and posture (orientation) of the individual workplaces are detected based on the captured image. For this purpose, images of a reference object, which is one of workplaces W, subjected to a picking-up operation or an object having a shape identical to that of the workpiece W are captured in different directions by the image capturing device and reference models are created from the image data obtained by the image capturing and stored in advance. Matching processing between the image data obtained by capturing the image of the pile of workplaces and the reference models is executed to select an image of one workpiece matched with one of reference models, and a position/posture of the selected workpiece is determined based on the selected image of the workpiece in the image field of view, the selected one of taught modes and the position/posture information being associated with the selected one of the reference models.

[0012] FIG. 3 is a block diagram showing a principal part of a robot controller 10 for use in the embodiment of the present invention. A main processor 1, a memory 2 including a RAM, a ROM and a nonvolatile memory (such as an EEPROM), an interface 3 for a teaching operating panel, an interface 6 for external devices, an interface 7 for an image processing apparatus and a servo control section 5 are connected to a bus 8. A teaching operating panel 4 is connected to the interface 3 for a teaching operating panel. ,

[0013] A system program for supporting basic functions of the robot RB and robot controller 10 are stored in the ROM of the memory 2. Robot operation programs and their related determined data which are taught in accordance with various operations are stored in the nonvolatile memory of the memory 2. The RAM of the memory 2 is used for temporary storage of data for various arithmetic operations performed by the processor 1.

[0014] The servo control section 5 comprises servo controllers 5al to 5an (n:sum of the number of all the axes of the robot including additional movable axes of a tool attached to a wrist of the robot), each composed of a processor, a ROM, a RAM, etc. Each servo controller performs position/velocity loop control and also current loop control for its associated servomotor for driving the axis, to function as a so-called digital servo controller for performing loop control of position, velocity and current by software. Each servomotor M1-Mn for driving each axis has its drive controlled according to outputs of the associated servo controller 5al-5an through the associated servo amplifier 5bl-5bn. Though not shown in FIG. 3, a position/velocity detector is attached to each servomotor MI-Mn, and the position and velocity of each servomotor detected by the associated position/velocity detector is fed back to the associated servo controller 5al-5an. Connected to the input/output interface 6 are sensors of the robot, and actuators and sensors of peripheral devices.

[0015] FIG 4 is a block diagram of the image processing apparatus 30 connected to an interface 7 of the robot controller 10. The image processing apparatus 30 comprises a processor 31 to which a ROM 32 for storing a system program to be executed by the processor 31, an image processor 33, an image-capturing- device interface 34 connected to the image capturing device 20, a MDI 35 with a display such as a CRT or a liquid crystal display for inputting and outputting various commands and data, a frame memory 36, a nonvolatile memory 37, a RAM 38 for temporary storage of data and a communication interface 39 for the robot controller are connected. An image captured by the camera 20 is stored in the frame memory 36. The image processor 33 performs image processing from images stored in the frame memory 36 on demand of the processor 31 so as to recognise an object. The architecture and function of the image processing apparatus 30 itself is no way different form the conventional image processing apparatus. The image processing apparatus 30 of the present invention is different form the conventional one in that reference models as described later are stored in the nonvolatile memory 37 and pattern matching processing is performed on an image of a pile of workpieces W captured by the image capturing device 20 using the reference models to obtain the position and posture of a workpiece W.

[0016] The image capturing device 20 is used for obtaining image data, as described later, and may be a CCD camera for obtaining two-dimensional images data or a visual sensor capable of obtaining three-dimensional image data including distance data. In the case of using the CCD camera, the image data is obtained by a conventional method based on two-dimensional images captured by the CCD camera, but in the case of the visual sensor capable of obtaining three-dimensional data including distance data, two-dimensional arrangement data with distance data between the sensor and an object is obtained. A visual sensor for obtaining the three-dimensional data including distance data is known, for example, from three-dimensional visual sensors of a spot light scanning type disclosed in Japanese Patent Publication No. 7-270137, and the summary of the three-dimensional visual sensor is described below.

[0017] This visual sensor detects a three-dimensional position of an object by irradiating a light beam to form a light spot on the object for scanning the object in two different directions (X direction and Y direction) and by detecting the light reflected on the object by a position sensitive detector (PSD). Three dimensional position of the object is measured by a calculation using the respective inclination angles $\theta x, \theta y$ of mirrors for scanning and an incident positions of the reflected light beam on the PSD.

[0018] Referring to FIGS. 7-9, a method of obtaining two-dimensional arrangement data including distance data using the three-dimensional visual sensor will be

explained briefly.

[0019] Scanning range (measuring range) on an object is set in advance, and an inclination angle θx, θy of the mirrors is controlled discretely. As shown in FIG. 7, the scanning is performed from a point (1,1) to a point (1, n), from a point (2,1) to a point (2, n), ..., from a point (m, 1) to a point (m, n) on the X-Y plane within the scanning range, to measure three-dimensional positions of each reflected point on the object. Also, a distance Z (i, j) between the sensor and the reflection point (i, j) on the object is obtained and stored in the RAM38 of the image processing apparatus 30. Thus, the image data is obtained as two-dimensional arrangement data including the distance data Z (i, j) between the sensor and the reflection point on the object, as shown in FIG. 8.

[0020] FIG. 9 is a flowchart of processing to be executed by the processor 31 of the image processing apparatus 30 for obtaining the image data.

[0021] First, indexes i and j are respectively set to "1" (Step 300) and the inclination angle (θx, θy) of the mirrors is set to (x1, yl) to direct to the start point (1, 1) and an irradiation command with the inclination angle is send to the sensor 20 (Steps 301-303). The sensor irradiates a light beam with the mirrors set at the inclination angle. The signal representing the image captured by the PSD is sent to the image processing apparatus 30. The processor 31 of the image processing apparatus 30 calculates the position of the reflection point on the object of the signal from the PSD and the inclination angle (θx, θy) of the mirrors to obtain the distance Z (i, j) between the sensor and the position of the reflection point on the object. This value Z (i, j) is stored in the RAM 38 as the two-dimensional arrangement data [i, j] (Step 304, 305). The calculation for obtaining the position of the reflection point and the distance Z (i, j) may be performed by the sensor 20.

[0022] Then, the index i is incrementally increased by "1" and the inclination angle θx of the mirror for X-axis direction scanning is increased by the predetermined amount Δx (Step 306, 307). It is determined whether or not the index i exceeds the set value n (Step 308). If the index i does not exceed the set value n, the procedure returns to Step 303 and the processing from Step 303 to Step 308 is executed to obtain the distance Z (i, j) of the next point. Subsequently, the processing of Steps 303-308 are repeatedly executed until the index i exceeds the set value n to obtain and store the distance Z (i, j) of the respective points (1, 1) to (1, n) shown in FIG. 7.

[0023] If it is determined that the index i exceeds the set value n in Step 308, the index i is set to "1" and the index j is incrementally increased by "1" to increase the inclination angle θy of the mirror for Y-axis direction scanning (Steps 309-311). Then, it is determined whether or not the index j exceeds the set value m (Step 312) and if the index j does not exceed the set value m, the procedure returns to Step 302 to repeatedly executes the processing of Step 302 and the subsequent Steps.

[0024] Thus, the processing from Step 302 to Step 312 is repeatedly executed until the index j exceeds the set value m. If the index j exceeds the set value m, the points in the measurement range (scanning range) shown in FIG. 7 have been measured entirely, the distance data Z (1, 1) - Z (m, n) as two dimensional arrangement data are stored in the RAM28 and the image data obtaining processing is terminated. A part of the image data of two dimensional arrangements or a plurality of distance data can be obtained by appropriately omitting the measurement of the distance for the index i.

[0025] The foregoing is a description on the processing for obtaining two dimensional arrangement data as image data using the visual sensor capable of measuring the distance. Using the two-dimensional arrangement data obtained in this way as image data, creation of reference models and detection of position and posture (orientation) of an object can be performed. In order to simplify the explanation, the following description will be made assuming that a CCD camera 20 is used as an image capturing device and the two dimensional image data obtained by capturing image of the object by this camera 20 is used.

[0026] Processing for creating reference models will be explained referring to FIGS. 2a-2d and FIG. 5. FIG. 5 is a flowchart showing processing for teaching reference models to the image processing apparatus 30 according to the present invention.

[0027] One reference workpiece (one of the workpieces W as object for robot operation or a workpiece having a three-dimensional shape identical to that of the workpiece W) is prepared for creating reference models. A first (0-th) position/posture of the reference workpiece at which the camera 20 attached to a distal end of a robot wrist captures the image of the object is set, and an axis of rotation and rotation angles with respect to the first (0-th) position/posture are set in order to determine the subsequent positions/postures of the reference workpiece. In addition, the number of positions/postures of the workpiece at which the camera 20 captures the image of the object is set. In this example, information of both position and posture is used. However it is sufficient for creating reference models to use only posture (orientation) information if the demanded precision of position is not high.

[0028] As shown in FIGS. 2a to 2d, in this example, images of the reference workpiece are captured from four different directions and reference models are created based on the four image data. As shown in FIG. 2a, an image of the reference workpiece is captured from the direction of a Z-axis of a world coordinate system at 0-th position/posture to create 0-th reference model. For setting the subsequent positions/postures, an axis perpendicular to an optical axis of the camera and passing a central point of the workpiece (origin of a work coordinate system set to the workpiece) and rotation angles of the workpiece along the rotation axis are set for this camera position. Since the optical axis of the

camera is set parallel to the Z axis of the world coordinate system, an axis parallel to either the X-axis or the Y-axis of the world coordinate system, which is perpendicular to the Z axis, can be selected and the workpiece is rotated around the rotation axis at the workpiece position.

[0029] In the example, an axis parallel to the X-axis of the world coordinate system is set as the rotation axis, and for the position/posture shown in FIG. 2b, the rotation angle of 30° is set to rotate the workpiece by 30° with respect to the camera along the rotation axis. A first reference model is created based on the image data of the workpiece at the position/posture shown in FIG. 2b. Similarly, as shown in FIGS. 2c and 2d, the workpiece is rotated by 60° and 90°, respectively, along the rotation axis for capturing images of the workpiece to create 2nd and 3rd reference models.

[0030] In this example, rotation angles of zero degree, 30 degrees, 60 degrees and 90 degrees are set for creating four reference models. The dividing range of the rotation angles may be set more finely and/or range of the rotation angle may be set greater to create more reference models for more precise detection of the position/posture of the workpiece.

[0031] The processing for creating the four reference models will be explained referring to flowchart of FIG. 5.

[0032] As described above, the 0-th position/posture of the robot at which the camera 20 captures the image of the object, and the rotation axis and the rotation angles with respect to the 0-th position/posture are set in advance in order to determine the subsequent positions/postures of the reference workpiece, and also the number of the subsequent positions/postures of the workpiece are set. For intelligible explanation, it is assumed that an optical axis of the camera is parallel to the Y-axis of the world coordinate system and that a position where the X-axis and Y-axis coordinate values are identical to those of the reference workpiece and only the Z-axis coordinate value is different from that of the position of the reference workpiece is taught to the robot as the 0-th image capturing position for obtaining the 0-th reference model. Further, the positions of the robot where the camera is rotated with respect to the reference workpiece by 30 degrees, 60 degrees and 90 degrees along the axis passing the central point of the reference workpiece and parallel to the X-axis of the world coordinate system are set as the 1st, 2nd and 3rd image capturing position, and the number N of the image capturing positions is set to "4."

[0033] When a command for creating reference models is inputted from teaching operation panel 4, the processor 1 of the robot controller 10 sets a counter M for counting the number of the image capturing to "0" (Step 100). The robot is operated to have the M-th position/posture and a command for image capturing is outputted to the image processing apparatus 30 (Step 101). In response to this command, the image processing apparatus 30 performs capturing of an image of the reference workpiece with the camera 20 and the captured image data is stored in the frame memory 36. Further, relative position/posture of the workpiece with respect to the camera is obtained and stored in the nonvolatile memory 37 as relative position/posture of M-th reference model, and a data-captured signal is sent to a robot controller (Step 103). Thus, position/posture of the workpiece in a camera coordinate system set to the camera is obtained from the position/posture of the camera and the position/posture of the reference workpiece in the world coordinate system when capturing the image by the camera, and is stored as the relative position/posture of the workpiece with respect to the camera. For example, the position/posture of the workpiece in the camera coordinate system is stored as [x0, y0, z0, α0, β0, γ0]c, where α, β and γ mean rotation angle around X-, Y-, Z- axes, and "c" means the camera coordinate system.

[0034] Upon receipt of the data-captured signal, the processor 1 of the robot controller 10 incrementally increases the value of the counter M by " 1 "(Step 104) and determines whether or not the value of the counter M is less than a set value N (=4) (Step 105). If the value of the counter M is less than the set value N, the procedure returns to Step 101 to move the robot to the M-th image-capturing position/posture. Thus, in the example as shown in FIGS. 2a-2d, the camera is successively turned by 30 degrees around the axis parallel to X axis of the world coordinate system and passing the workpiece position, and successively captures the image of the workpiece, and reference models and relativity positions/postures of the camera with respect to the workpiece at the image capturing are stored.

[0035] Processing of Steps 101-105 is repeatedly executed until the value of the counter M equals to the set value N (=4), and the reference models and the relative positions/postures of the camera and the workpiece are stored in the nonvolatile memory 37. Thus, the reference models created from the image data of the workpiece at the positions/postures shown in FIGS. 2a-2d are stored, and the relative positions/postures between the camera and the workpiece for respective reference models are stored as positions/postures of the workpiece W in the camera coordinate system as [x 0, y 0, z 0, α0, β0, γ0]c,[x 1, y 1, z 1, α1, β1, γ1]c, [x 2, y 2, z 2, α2, β2, γ2]c, and [x 3, y 3, z 3, α3, β3, γ3]c.

[0036] The reference models and the relative position/posture of the workpiece W and the camera 20 are stored in the nonvolatile memory 37 of the image processing apparatus 30. In the above described embodiment, the reference models are created using a robot, however, the reference models may be created by a manual operation without using a robot. In this case, the reference workpiece is arranged within the field of view of the camera connected to the image processing apparatus 30, and the images of the workpiece with different postures are captured by the camera. The reference models are created based on the image data and

the relative positions/postures of the camera and the workpiece at the image capturing manually inputted, and are stored with the respective relative positions/ postures.

[0037] The reference models may be created from a part of the image data of the reference object, and may be created by processing the image data of the reference object.

[0038] In addition, the reference models may be created based on the stored image data of the reference workpiece when detecting the position/posture of the objective workpiece, without creating and storing the reference models in advance.

[0039] Hereinafter, a picking-up operation for taking out an individual workpiece by a robot from a pile of workpieces each having a shape identical to that of the reference workpiece will be described, as an example of a method of detecting three-dimensional position/ posture of an object, using the image processing apparatus 30 storing the reference models.

[0040] FIG. 6 is a flowchart of the carrying out of the picking-up operation. When a picking-up command is inputted into the robot controller 10 from the teaching operation panel 4, the processor 1 operates the robot RB to move the camera attached to the robot wrist to an image capturing position where a pile of workpieces is within a field of view of the camera 20 (Step 200). The three dimensional position/posture of the camera 20 on the world coordinate system at this image capturing position is outputted to the image processing apparatus 30, and an image capturing command is outputted (Step 201). Upon receipt of the image capturing command, the processor 31 of the image processing apparatus 30 captures an image of the pile of the workpieces W. to obtain image data of some workpieces W and store the data in the frame memory 36 (Step 202).

[0041] Then, pattern matching processing is performed for the image data stored in the frame memory 36 using one of the reference models (the first reference model) stored in the nonvolatile memory 37 so as to detect a workpiece W (Step 203). In this pattern matching processing, matching of the image data of the reference model with the image data of workpieces is performed on the basis of position, turn and scale. It is determined whether or not an object has a matching value equal or greater than the set value (Step 204). If an object having a matching value equal or greater than the set value is not detected, the procedure proceeds to Step 205 to determine whether or not the pattern matching is performed using all the reference models (1st to 4th reference models). If the pattern matching using all the reference models is not yet performed, further pattern matching is performed using another reference model (Step 206).

[0042] If it is determined in Step 204 that an object having a matching value equal or greater than the set value with respect to any of the reference models is detected, the procedure proceeds to Step 207 to perform matching processing on the two-dimensional data of the detected workpieces W using every taught mode. In Step 208, the reference model having the largest matching value in the pattern matching processing is selected, and the relative position/posture of the workpiece W with respect to the camera 20 is determined based on the relative position/posture of the camera and the reference workpiece stored for the selected reference model, and position, rotation angle and scale of the image of the workpiece in the matching processing, (Step 208). The position and posture (orientation) of the detected workpiece on the world coordinate system is determined from the position and posture of the camera 20 in the world coordinate system, which has been sent in Step 201, and the relative position/posture of the workpiece W with respect to the camera 20, and is outputted (Step 209). Thus, since the relative position/posture of the workpiece W with respect to the camera 20 is the position/posture of the workpiece W in the camera coordinate system, the position and posture (orientation) of the detected workpiece W in the world coordinate system is obtained by an arithmetic operation of coordinate transformation using the data of the position/posture of the workpiece W in the camera coordinate system and the position/posture of the camera 20 in the world coordinate system (Step 209).

[0043] The reference model having the highest matching value is selected in this embodiment, although a reference model of the rotation angle of zero degree (the O-th reference model) may be selected in accordance with precedents, or an object having the highest expansion rate of scale (the object which is nearest to the camera, i.e. located at the suit of the pile in this example) may be selected in accordance with precedents.

[0044] The robot controller 10 operates the robot to perform a picking-up operation to grip and hold the detected workpiece W and move the held workpiece W to a predetermined position, based on the three-dimensional position/posture of the workpiece W (Step 210). Then, the procedure returns to Step 202 to repeatedly execute the processing of Step 202 and subsequent Steps.

[0045] When all the workpieces have been picked-up from the pile of the workpieces, a matching value equal to or greater than the set reference value cannot be obtained in the pattern matching processing for all reference models in Steps 203-206, and the picking-up operation is terminated.

[0046] In the case where a pile of the workpieces cannot fall within the field of view of the camera 20, or in the case where it is not necessary to capture an image of a workpiece behind other workpieces by changing the orientation of the camera, the procedure may return to Step 200 when "Yes" is determined in Step 205, to move the camera to another position/posture at which an image of the object workpiece can be captured.

[0047] In addition, in the case where the robot and the image processing apparatus 30 are used in combination

as in the foregoing embodiment, the robot controller 10 may store the three-dimensional position/posture of the camera without outputting it to the image processing apparatus 30 in Step 201 and the relative position/posture of the workpiece and the camera may be outputted from the image processing apparatus 30 to the robot controller 10 in Step 208 to execute the processing of Step 209 in the robot controller 10.

[0048] Further, in the case where a wide-angle lens is installed in the CCD camera as the image capturing device, for example, there is possibility of judging the inclination angle to be 30 degrees by influence of parallax when a workpiece of zero degree inclination is at a corner of a field of view of the camera. In such a case, the camera may be moved parallelly in accordance with the position of the workpiece in the field of view of the camera to a position right above the workpiece so that the effect of parallax is lost, and at this position the image capturing processing of Step 201 and the subsequent Steps in FIG. 6 is performed so that a false judgment is prevented.

[0049] Furthermore, in order to obtain three-dimensional position/posture of an object workpiece whose three-dimensional position/posture is unknown without using a robot, the camera is arranged to capture an image of a pile of workpieces or a region containing the objective workpiece within a field of view of the camera, and the position/posture of the camera in the world coordinate system is inputted to the image processing apparatus 30 and an object detection command is issued to the image processing apparatus 30, to make the image processing apparatus 30 execute Steps 202-209 of FIG. 6.

[0050] The image data for creating the reference models may be obtained at a place different form the place where the robot is installed. In this case, the image data may be supplied to the image processing apparatus on line through a communication interface provided in the image processing apparatus, or may be supplied offline through a disc driver for reading a floppy disk, etc.

[0051] According to the present invention, the position/posture of an object workpiece in a randomly arranged pile of workpieces or an aggregation of workpieces gathered in a predetermined region which have identical shapes and different three-dimensional positions/postures is detected, to thereby enable a robot to automatically pick out an individual workpiece from such a pile or an aggregation.

Claims

1. An image processing apparatus for detecting posture, or posture and position, of an object comprising:

   an image capturing device;
   a memory storage reference model created based on image data of a reference object captured by said image capturing device in a plurality of directions, and storing information of the capturing directions to be respectively associated with said reference models, said reference object being the object for detection or an object having a shape identical to that of the object for detection; and
   a processor to perform matching processing of image data containing an image of the object for detection captured by said image capturing device with image data for said reference models to select an image of an object matched with one of said reference models, and to obtain posture, or posture and position, of the object based on the selected image of the object, said one reference model and the information of the direction associated with said one reference model.

2. An image processing-apparatus according to claim 1, wherein said reference models comprises a part of the image data of the reference object.

3. An image processing apparatus according to claim 1 or 2, wherein said reference models are obtained by processing the image data of the reference object.

4. An image processing apparatus according to any preceding claim, wherein said image capturing device comprises a camera for capturing two-dimensional image data.

5. An image processing apparatus according to claim 4, wherein said image data of the reference object are captured by said image capturing device from a predetermined distance.

6. An image processing apparatus according to any one of claims 1 to 3, wherein said image capturing device comprises a visual sensor for capturing three-dimensional image data.

7. An image processing apparatus according to claim 6, wherein said image data containing an image of the object for detection captured by said visual sensor are two-dimensional arrangement data including distance information from the object of detection to the image capturing device, a part of said two-dimensional arrangement data or a set of distance data.

8. An image processing apparatus according to an one of claims 1 through 7, wherein said image capturing device is attached to a robot.

9. An image processing apparatus according to claim

1, wherein said image data of the reference object are captured in a place different from a place where the detection of the object is performed, and supplied to the image processing apparatus on line or off line.

FIG. 1

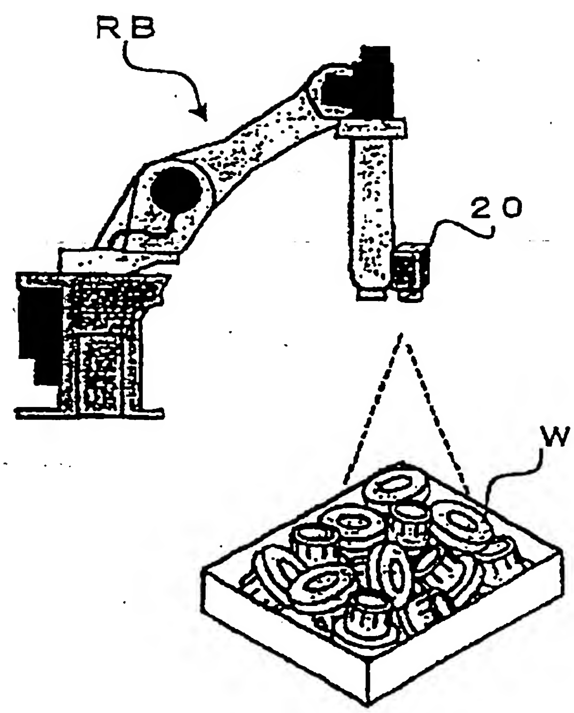


FIG. 2a

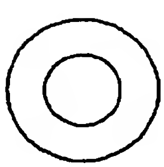


FIG. 2b

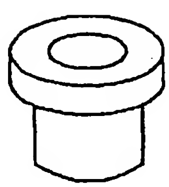


FIG. 2c

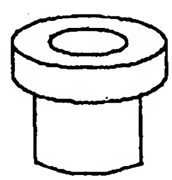


FIG. 2d

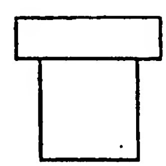


FIG.3

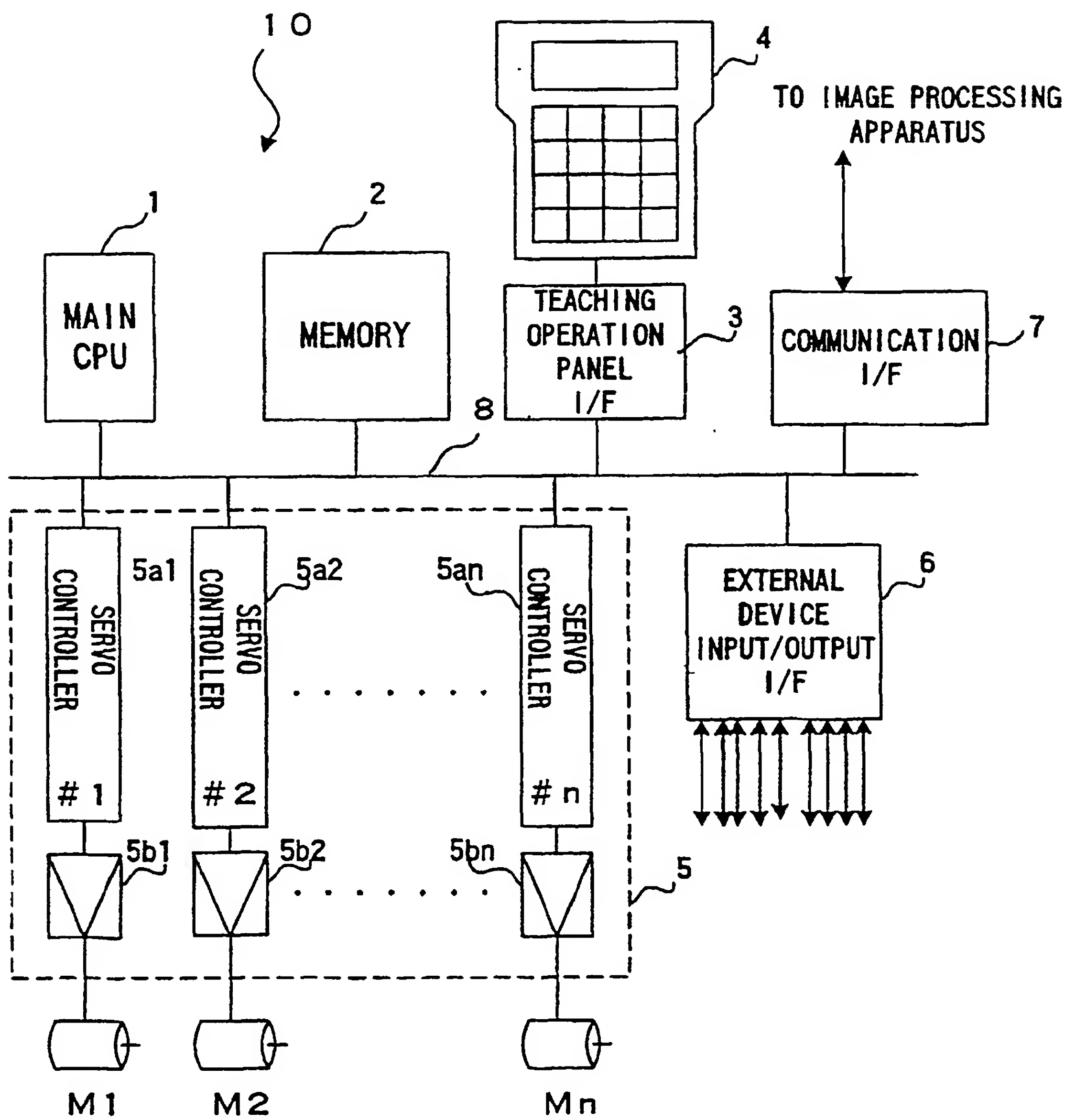

10
4
TO IMAGE PROCESSING APPARATUS
1
2
MAIN CPU
MEMORY
TEACHING OPERATION PANEL I/F
3
COMMUNICATION I/F
7
8
5a1
5a2
5an
SERVO CONTROLLER #1
SERVO CONTROLLER #2
SERVO CONTROLLER #n
EXTERNAL DEVICE INPUT/OUTPUT I/F
6
5b1
5b2
5bn
5
M1
M2
Mn

FIG. 4

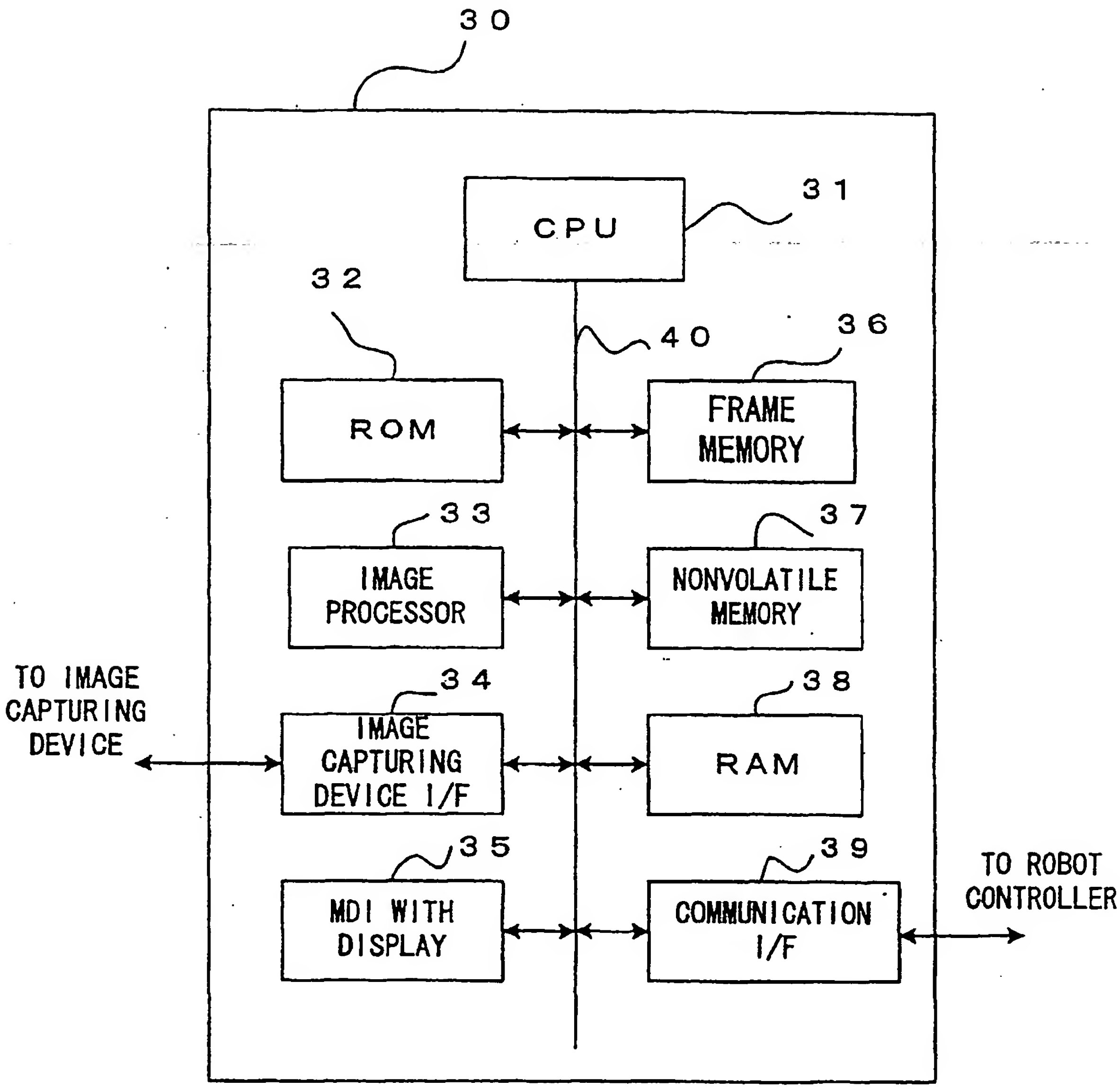

30
CPU
31
32
40
36
ROM
FRAME MEMORY
33
37
IMAGE PROCESSOR
NONVOLATILE MEMORY
TO IMAGE CAPTURING DEVICE
34
38
IMAGE CAPTURING DEVICE I/F
RAM
35
39
TO ROBOT CONTROLLER
MDI WITH DISPLAY
COMMUNICATION I/F

FIG. 5

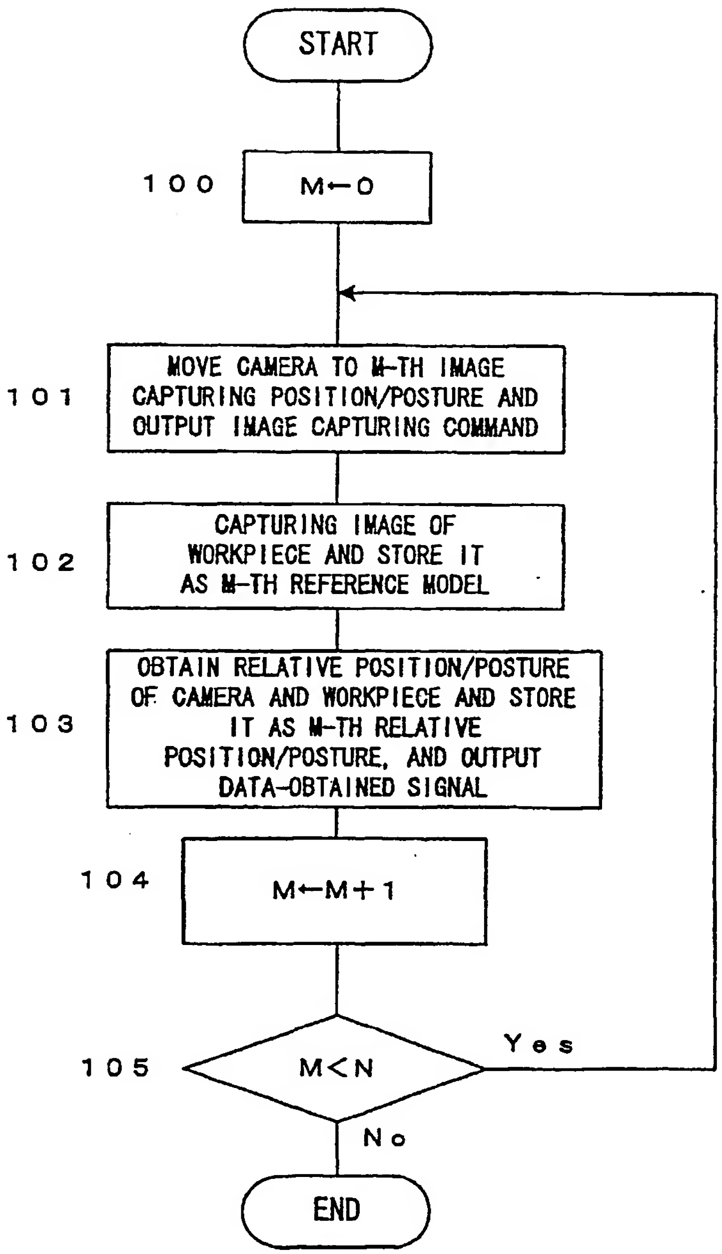

START
100
M←0
101
MOVE CAMERA TO M-TH IMAGE CAPTURING POSITION/POSTURE AND OUTPUT IMAGE CAPTURING COMMAND
102
CAPTURING IMAGE OF WORKPIECE AND STORE IT AS M-TH REFERENCE MODEL
103
OBTAIN RELATIVE POSITION/POSTURE OF CAMERA AND WORKPIECE AND STORE IT AS M-TH RELATIVE POSITION/POSTURE, AND OUTPUT DATA-OBTAINED SIGNAL
104
M←M+1
105
M<N
Yes
No
END

FIG. 6

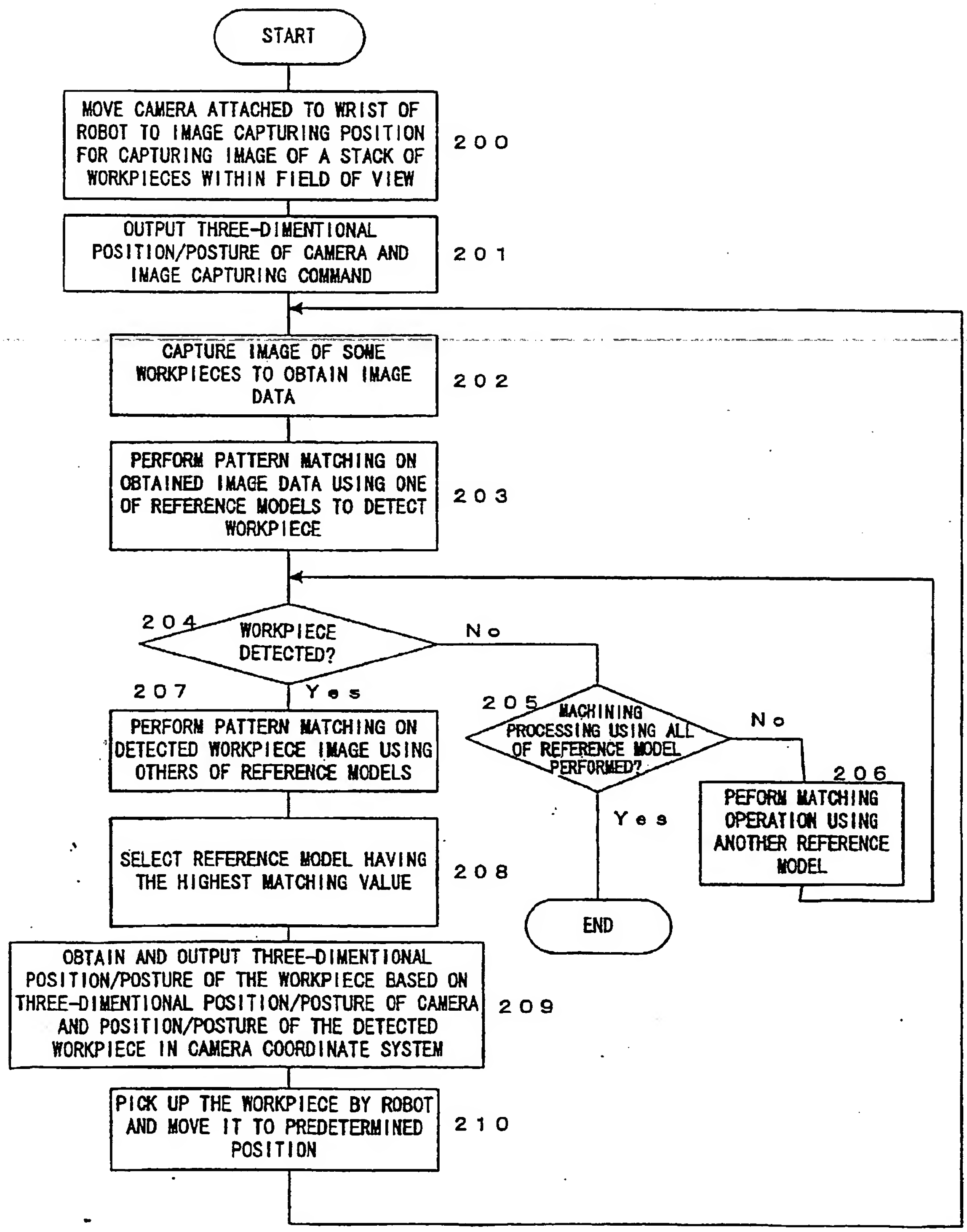

START
MOVE CAMERA ATTACHED TO WRIST OF ROBOT TO IMAGE CAPTURING POSITION FOR CAPTURING IMAGE OF A STACK OF WORKPIECES WITHIN FIELD OF VIEW
200
OUTPUT THREE-DIMENTIONAL POSITION/POSTURE OF CAMERA AND IMAGE CAPTURING COMMAND
201
CAPTURE IMAGE OF SOME WORKPIECES TO OBTAIN IMAGE DATA
202
PERFORM PATTERN MATCHING ON OBTAINED IMAGE DATA USING ONE OF REFERENCE MODELS TO DETECT WORKPIECE
203
204
WORKPIECE DETECTED?
No
Yes
205
MACHINING PROCESSING USING ALL OF REFERENCE MODEL PERFORMED?
No
Yes
206
PEFORM MATCHING OPERATION USING ANOTHER REFERENCE MODEL
END
207
PERFORM PATTERN MATCHING ON DETECTED WORKPIECE IMAGE USING OTHERS OF REFERENCE MODELS
SELECT REFERENCE MODEL HAVING THE HIGHEST MATCHING VALUE
208
OBTAIN AND OUTPUT THREE-DIMENTIONAL POSITION/POSTURE OF THE WORKPIECE BASED ON THREE-DIMENTIONAL POSITION/POSTURE OF CAMERA AND POSITION/POSTURE OF THE DETECTED WORKPIECE IN CAMERA COORDINATE SYSTEM
209
PICK UP THE WORKPIECE BY ROBOT AND MOVE IT TO PREDETERMINED POSITION
210

FIG. 7

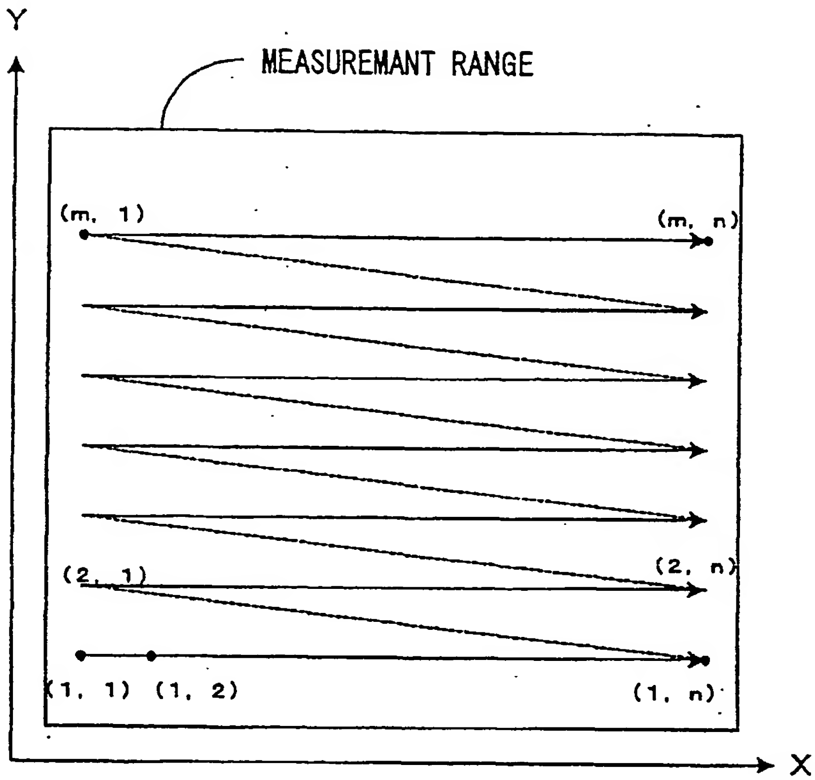
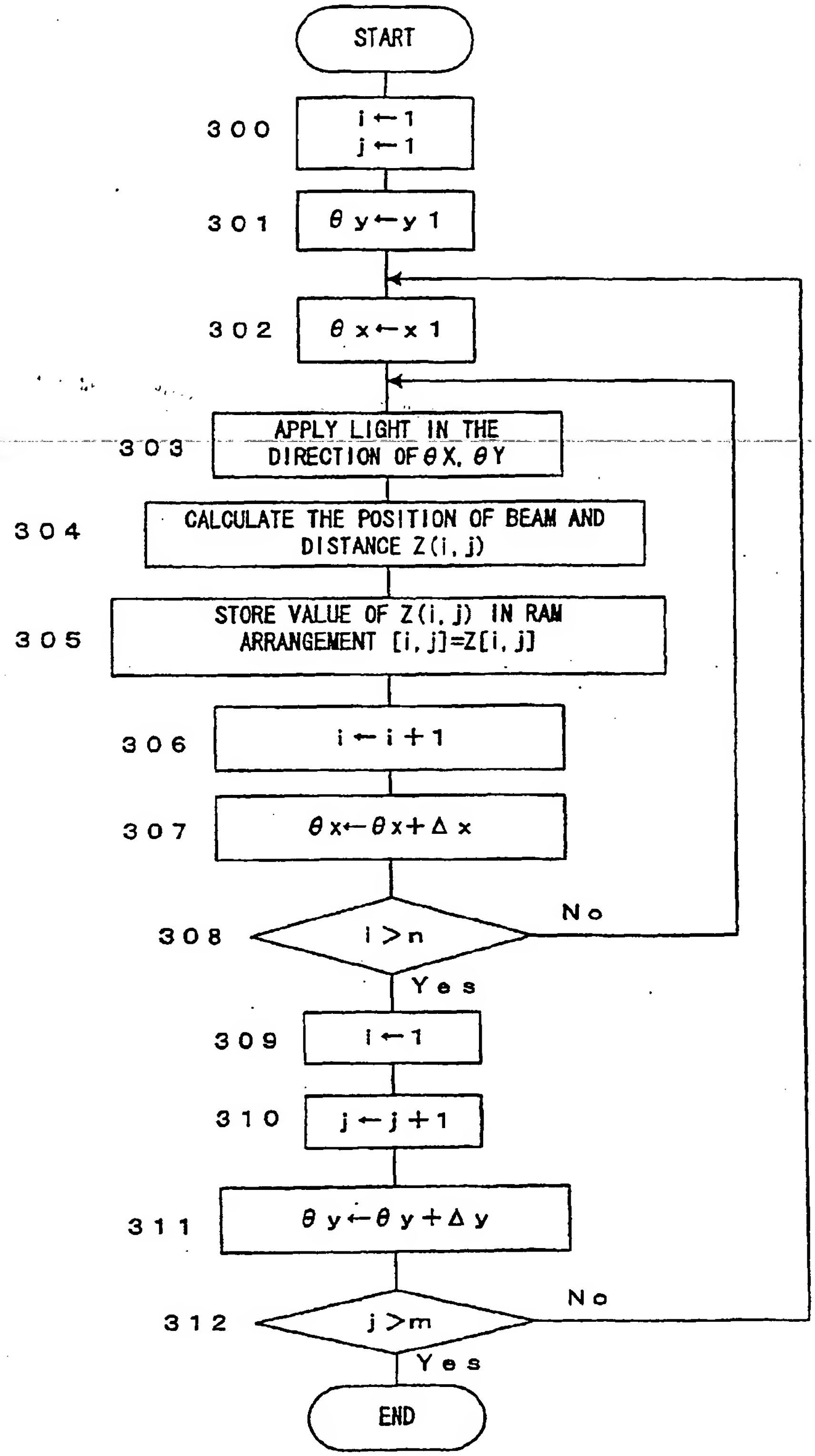


FIG. 8

Z(m, 1)	· · ·	· · ·	· · ·	· · ·	Z(m, n)
· · ·	· · ·	· · ·	· · ·	· · ·	· · ·
· · ·	· · ·	· · ·	· · ·	· · ·	· · ·
· · ·	· · ·	Z(i, j)	· · ·	· · ·	· · ·
· · ·	· · ·	· · ·	· · ·	· · ·	· · ·
Z(1, 1)	Z(1, 2)	· · ·	· · ·	· · ·	Z(1, n)

FIG. 9


START
300
i ← 1
j ← 1
301
θ y ← y 1
302
θ x ← x 1
303
APPLY LIGHT IN THE DIRECTION OF θX, θY
304
CALCULATE THE POSITION OF BEAM AND DISTANCE Z(i, j)
305
STORE VALUE OF Z(i, j) IN RAM ARRANGEMENT [i, j]=Z[i, j]
306
i ← i + 1
307
θ x ← θ x + Δ x
308
i > n
No
Yes
309
i ← 1
310
j ← j + 1
311
θ y ← θ y + Δ y
312
j > m
No
Yes
END